AUGMENTING INDIANA'S GROUNDWATER LEVEL MONITORING NETWORK: OPTIMAL SITING OF ADDITIONAL WELLS TO ADDRESS SPATIAL AND CATEGORICAL SAMPLING GAPS

Benjamin J. Sperl

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Science
in the Department of Geography,
Indiana University

December 2014

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Master's Thesis Committee

Aniruddha Banerjee, Ph.D., Chair

Vijay Lulla, Ph.D.

Frederick Bein, Ph.D.

ACKNOWLEDGEMENTS

I must thank my committee members, professors Rudy Banerjee, Vijay Lulla, and Rick Bein, for their technical guidance, keen observations, and for taking me deeper into the fields of geography and location science. I express sincere gratitude to Randy Bayless, Ph.D., who has been my principal tutor in hydrology and an outstanding professional mentor. This project began in his classroom and his input has been helpful at every juncture. Particular thanks are extended to the world-class leadership and scientists at the Indiana-Kentucky Water Science Center for supporting me in my research and clarifying my vision as to where I'd like to go with my career. I would also like to thank colleagues from the Indiana Geological Survey, The Polis Center, and the Indiana Department of Transportation for everything they've taught me over the past couple of years. I subconsciously use the skills and knowledge imparted to me by these people on a regular basis. Most importantly, I'm grateful for my family members who have motivated, inspired, and supported me in every way possible. I could not be where I am now without their help.

Benjamin J. Sperl

AUGMENTING INDIANA'S GROUNDWATER LEVEL MONITORING NETWORK: OPTIMAL SITING OF ADDITIONAL WELLS TO ADDRESS SPATIAL AND CATEGORICAL SAMPLING GAPS

Groundwater monitoring networks are subject to change by budgetary actions and stakeholder initiatives that result in wells being abandoned or added. A strategy for network design is presented that addresses the latter situation. It was developed in response to consensus in the state of Indiana that additional monitoring wells are needed to effectively characterize water availability in aquifer systems throughout the state. The strategic methodology has two primary objectives that guide decision making for new installations: (1) purposive sampling of a diversity of environmental variables having relevance to groundwater recharge, and (2) spatial optimization by means of maximizing geographic distances that separate monitoring wells. Design objectives are integrated in a discrete facility location model known as the p-median problem, and solved to optimality using a mathematical programming package.

Aniruddha Banerjee, Ph.D., Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

At present, the U.S. Geological Survey (USGS) and Indiana Department of Natural Resources (IDNR) maintain an active network of 37 wells equipped with hardware to record and transmit water levels continuously. Data from these wells are steadily building a historical base for detecting seasonal and long-term trends, and for observing how groundwater levels respond under a variety of conditions. As demand for groundwater increases and water levels decline, the imperative grows for a more thorough understanding of available storage in the state's aquifer systems (Wittman 2014). Hence, a strategy is needed that will optimize public investment in an expansion of Indiana's groundwater level monitoring network.

BACKGROUND

Groundwater monitoring programs are implemented with the overarching objective to improve understanding of aquifer conditions and behavior. Beyond this common purpose, a multitude of factors elicit contrast among individual programs with specific design considerations. Planning for a network of monitoring wells may be complicated by elements of scale (e.g., local, state, regional), parameters sought (e.g., water quality vs. water levels), vertical dimension (i.e., monitoring depths), and specific statements of exactly what types of interactions network wells are intended to observe (e.g., response to anthropogenic stress, groundwater surface-water interactions). Additionally, design problems may be further classified according to whether a network is being augmented, reduced, reconfigured, or starting anew.

An extensive body of research exists that covers a variety of approaches for the design and planning of groundwater monitoring programs. Heath (1976) outlines three sub-objectives of a monitoring network: (1) depict the areal extent of aquifers and changes in the potentiometric surface, (2) observe how aquifers respond to drawdown from pumping facilities, and (3) monitor “baseline” conditions where groundwater levels are unaffected by pumping (Heath 1976). Hudak and Loaiciga (1992) solve the maximal covering location problem for adding new wells to a network intended to maximize detection of a contaminant plume in a buried valley aquifer (Hudak, Loaiciga 1992; Church, ReVelle 1974). Fisher (2013) developed a tool that identifies a pre-specified number of wells whose removal from the network causes the smallest increase in variance associated with kriging-based interpolations of the water table. Pearson, Falteisek, and Berg (2011) designed a framework for the State of Minnesota that determines appropriate well densities for aquifers based on intensity of groundwater use. Nabi, Gallardo, and Ahmed (2011) applied principal component analysis and kriging to identify monitoring wells contributing redundant information in the spatial distribution of water quality parameters throughout a watershed. Zhou et al. (2002) conducted a groundwater tracing study in a karst environment of southern Indiana to locate monitoring wells near springs found to be in hydrologic connection with a landfill site.

The problem presented in this study involves adding new wells to an existing statewide groundwater level monitoring network with the objective that new monitoring sites will simultaneously address spatial and categorical sampling gaps. The regionally constrained p-median problem presented by Church (1990) is used to generate a spatially optimal distribution of new wells within the constraint that categorical sampling gaps known *a priori* are satisfied. An approach is taken that treats the network design as a discrete facility location model, which places this study under the broad field of operations research. Optimization problems in spatial analysis that cannot be solved using analytical methods such as kriging statistics may find advantages in the algorithmic (trial and error) methods offered by operations research.

DATA

Several layers of spatial information were necessary to optimally locate new monitoring wells. All data used in this project were obtained from online sources with the exception of some information that was procured by personnel at the U.S. Geological Survey. Spatially referenced data sets of significant water withdrawal facilities (SWWFs), locations of public facilities (e.g., recreation areas, academic institutions), land cover, elevation, cadastral boundaries, and aquifer system delineations were downloaded without charge from web-based data portals and government clearinghouses. While some generalization is inherent to the original data sets, the level of detail that they provide is considered adequate for the sampling requirements of this network design. However, some data sources offered more detail than desired, which required binning of continuous raster values into nominal categories and condensing enumerated attribute values into fewer classifications. Process steps for reformatting and deriving information from source data are explained with more detail in the sections that follow.

Aquifer Delineations

Two ESRI polygon shapefiles named "Aquifer_Systems_Unconsolidated_IDNR_IN.shp" and "Aquifer_Systems_Bedrock_IDNR_IN.shp" were downloaded from Indiana's open geospatial data clearinghouse known as IndianaMAP (IDNR 2011). Data from these files consist of polygons depicting the boundaries of unconsolidated and bedrock aquifer systems, and textual descriptions of their physical properties, which are assumed to be relatively homogenous. These delineations provide a basis for defining Indiana's aquifers as discrete sampling units in the network design.

NWS Climate Divisions

Boundaries of Indiana counties were downloaded as a shapefile from the U.S. Census Bureau TIGER/Line® webpage (Commerce 2014). A new field was added to the attribute table and records (for each county) were populated with the names of National Weather Service (NWS) climate divisions. The "Dissolve" tool of ESRI ArcToolbox was subsequently used to produce a data set of Indiana's nine climate divisions.

Topography

Digital elevation models (DEMs) with cell resolutions of one arc second (approximately 30 meters) were downloaded from The National Map for each of the 1x1 degree tiles intersecting Indiana (USGS 1999). Tiles were mosaicked into a single DEM and a slope (%) grid was derived using a 3x3 moving kernel (“Slope” tool of ArcToolbox). Continuous slope values were reclassified into the following bins: 0-12%, 13-23%, and 24-100%, and assigned nominal values of “Low”, “Moderate”, and “High”, respectively.

Land Cover

The 2006 version of the National Land Cover Dataset (NLCD) was downloaded from IndianaMAP as a TIFF file with 30-meter resolution (USGS 2006). The fifteen land cover classifications contained in the NLCD were reduced to a shorter list as follows: “Developed, Open Space”, “Developed, Low Intensity”, “Developed, Medium Intensity”, and “Developed, High Intensity” were condensed to a single classification of “Developed”; “Deciduous Forest”, “Mixed Forest”, and “Evergreen Forest” were condensed to “Forest”; and “Emergent Herbaceous Wetlands” and “Woody Wetlands” were condensed to “Wetlands.” A 500-meter majority filter (“Majority Filter” tool from ArcToolbox) was then applied to the grid so that values of predominant land cover types (within buffer area) were assigned to grid cells. Rationale for smoothing the grid was to simplify land cover heterogeneity into single classifications that are assumed to have the most dominant influence on local groundwater recharge (Nolan 2001). A radius of 500 meters was chosen because it has been applied in numerous other studies that assess the effects of land cover/land use on recharge and contamination to shallow groundwater wells (Nolan 2001). Many of these studies are in connection with the National Water-Quality Assessment (NAWQA) Program administered by the USGS.

Potential Monitoring Sites

A data set of potential sites for new monitoring wells was compiled from multiple shapefiles depicting locations for (1) wells in other networks (e.g., NAWQA), (2) publicly managed parks and recreation areas, (3) police and fire stations, (4) academic institutions, and (5) religious facilities. Attribute tables were reformatted to an agreeable standard and points were merged into a single shapefile.

Significant Water Withdrawal Facilities (2012)

Significant water withdrawal facilities are defined as any facility having the capacity to pump more than 100,000 gallons of water per day (IDNR 2014). A database of registered SWWFs is maintained by the State of Indiana and made publicly available in a downloadable spreadsheet format from the IDNR website. The database includes information on water use, monthly withdrawals, pumping capacity, and latitude/longitude coordinates of facilities. Locations of SWWFS were converted to shapefile format using the coordinates provided in the spreadsheet.

HARDWARE/SOFTWARE

Processing and analysis for this study was performed on a PC built with an Intel® Core™ i7 processor (2.80 Ghz) and 8.00 GB of installed memory (RAM). Three software programs were used: Microsoft Excel 2010; ESRI ArcGIS for Desktop version 10.1, which was used for formatting, cartographic output, and some processing tasks; and R (Project for Statistical Computing) version 3.1.0, which was used for processing and output of analytical results.

METHODOLOGY

Acknowledging that (1) hydraulic gradients and groundwater levels vary at multiple spatial and temporal scales, and (2) that the full complexity of the potentiometric surface is prohibitively difficult to depict at the scale of a statewide monitoring network, the methodology described herein attempts to distribute monitoring wells over a set of unique aquifer scenarios such that a wide range of environmental variables are represented by the network configuration. A design strategy is proposed that augments the existing IDNR-USGS network in such a way that addresses categorical sampling gaps, while maximizing geographic distances between proposed monitoring sites. The strategy can be conceptualized into three phases: (1) stratifying the state into smaller sampling units, (2) identifying spatial and categorical sampling gaps, and (3) solving for an optimal network configuration with maximal inter-well spacing.

Stratification by Spatial Intersections

Stratification of Indiana was performed via a union intersection of the domain values from four data sets representing variables of interest to the network design. Data sets of land cover, topography, climate, and aquifer systems were used as input to the union operation. Combinations of variables formed by the union operation are perceived to cover the full set of unique aquifer scenarios throughout Indiana. In this context, an aquifer scenario refers to a complete environment comprised of the aquifer material itself, local climate, and characteristics of the overlying ground surface (e.g., land cover and topography). Aquifer scenarios were assigned index values that are referenced in tables and figures throughout this report.

Categorical Sampling Gaps

The statewide network is intended to characterize Indiana's groundwater resource in terms of its availability and the variables that influence recharge. Therefore, a metric is needed that grades the network based on a quantity of the resource that it has characterized, or alternatively, how much remains to be explained. Such a metric might be presented as a ratio or percentage, which would indicate the amount that has been explained, or “covered”, in relation to the total. The question then becomes one of how to measure or represent the quantity upon which “coverage” is being evaluated. In this case, quantities of areal extent (square miles) and annual withdrawals from SWWFs (millions of gallons) were summarized for each of the indexed aquifer scenarios. Summarizing these quantities provides a means for ordering the aquifer scenarios by their importance in terms of groundwater usage and geographic pervasiveness. Categorical sampling gaps can then be determined as those aquifer scenarios having high importance, but lacking monitoring wells.

Aquifer scenarios with broader areal extent are placed at a higher priority for monitoring (annual withdrawals are only presented as a secondary measure) so that categorical coverage is maximized with each new well installation. This metric of (categorical) coverage implies that the network requires one or more wells per aquifer scenario to effectively observe the interrelationships among environmental variables and their influence on groundwater levels. Coverage in this context does not assert that monitoring wells provide accurate predictions of actual water levels in all areas having similar environments.

An additional stipulation for categorical coverage is now introduced. Wells from the active network that experience intermittent drawdown from nearby pumping facilities are excluded from the definition of coverage. In other words, bias from water level fluctuations induced by pumping must be negligible or absent altogether as a requirement that coverage be claimed in this analysis. This condition was set because natural interactions between groundwater levels and other aspects of the environment are best examined in isolation from a pumping variable (Heath 1976). Wells in the active network have been labeled "affected" or "unaffected" based on written observations noted during visits to wells by USGS hydrologic technicians. Of the 37 total wells in the active network, 21 have been labeled "unaffected" or "baseline" wells, whereas the remaining 16 are considered affected.

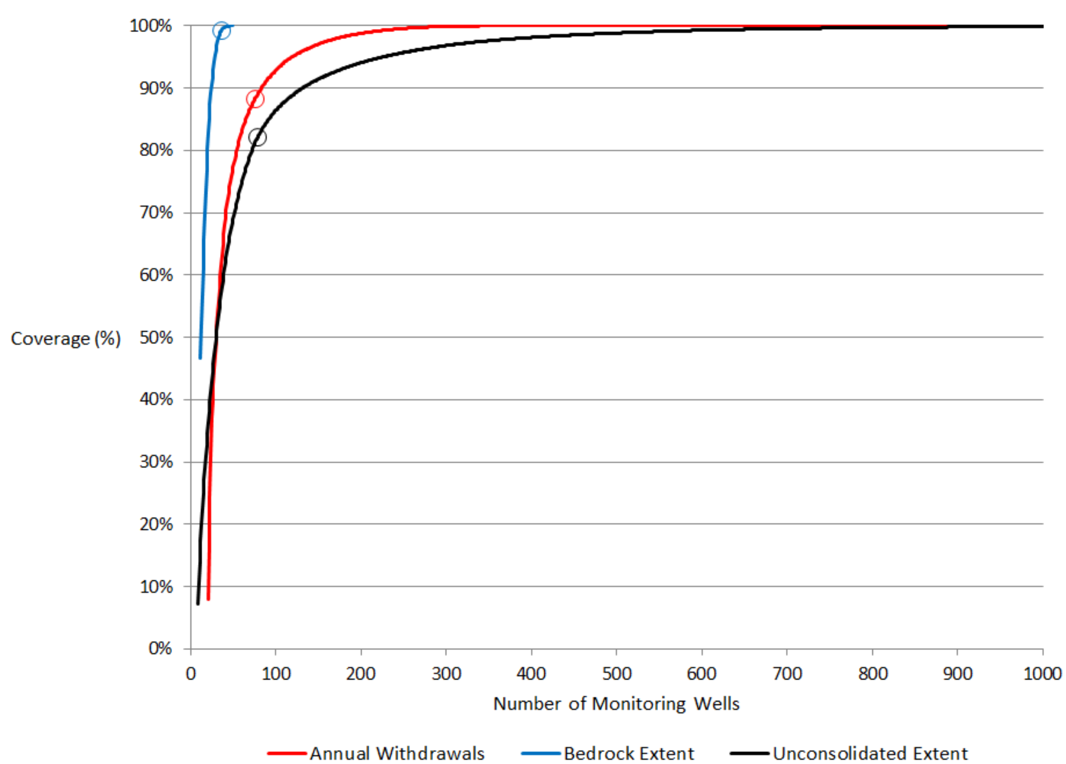
Prioritized Sampling

The number of wells needed to fully characterize Indiana’s groundwater resource may vary as a function of numerous factors, such as the granularity with which variables are represented or definitions of coverage. While it may seem reasonable to assert that the appropriate number of monitoring wells is closely linked to the number that it takes to achieve a goal of 100 percent coverage, the caveat to this is that aquifer scenarios are weighted according to their areal extent. Thus, a one-to-one relationship does not exist where every new well installation translates to an equal increment of improved network value.

Figure 1 resembles what is commonly referred to as a "return on investment" (ROI) curve. The ROI curve shows cumulative gain of coverage as hypothetical wells are added to the network. The three lines on the graph represent different quantities (i.e., annual withdrawals and areal extent of bedrock/unconsolidated aquifer scenarios) upon which coverage is measured. There are two separate lines for unconsolidated and bedrock aquifer scenarios because this analysis treats those settings as separate and parallel surfaces. It should also be noted that these lines were plotted independent of one another, which means the sequences for selecting aquifer scenarios in each line are different.

Points on the graph are plotted in a sequence that assigns wells to aquifer scenarios in order of their value (i.e., how much they improve cumulative coverage). Initial points on the graph represent current coverage by the active network of baseline wells. Currently, there are 9 wells monitoring 9 unique unconsolidated aquifer scenarios with 7.2% coverage; 12 bedrock wells monitoring 9 unique bedrock aquifer scenarios with 46.7% coverage; and combined the 21 baseline wells are covering 7.9% of total withdrawals. In the graph it is evident that early installations yield more substantial improvements in coverage and network value, and wells added later in the sequence have progressively smaller contributions to overall coverage.

Figure 1. Graph of coverage (%) vs. new well installations



The role of the ROI graph in this analysis was to facilitate a decision as to which aquifer scenarios should be prioritized to receive new wells. Visual inspection of the graph led to a conclusion that diminishing returns begin to set in when new wells start adding less than 0.25% coverage to the cumulative total. These threshold points are demarcated with hollow circles on the plotted lines. For annual withdrawals this occurs after 54 new wells, for bedrock extent it occurs after 24, and for unconsolidated extent it occurs after 70. Three separate lists of sampling needs were produced by selecting aquifer scenarios beneath the cutoff values for each line on the ROI graph. The lists were then combined in a full outer join to eliminate duplicates and guarantee that no records were excluded. The result of the join is a combined table of 142 total aquifer scenarios, which includes those already having baseline coverage as well as the unsampled scenarios that may be targeted to receive additional monitoring wells (Table 7).

Spatial Sampling Gaps

Evaluating the network purely on its coverage of qualitatively defined aquifer scenarios would neglect to address gaps that occur spatially. A definition for spatial coverage is needed that can be used to measure the extent to which areas are underserved by monitoring wells. Heath (1976) recommends a range of well densities for regional monitoring of groundwater levels:

> "The density of wells – that is, the number of wells per unit of area – needed in a hydrologic network would depend on the complexity of the system and the level of detail desired. It may range from more than 100/1,000 mi^2 for a complex area to be mapped in considerable detail to 2/1,000 mi^2 for a large area in which only the major features are mapped." (Heath 1976)

The upper bound of this range (1 well per 500 mi^2) is adopted by this study to evaluate spatial coverage. The radius of a circle with an area of 500 mi^2 is 12.62 miles. This value was used as an input parameter to the "Buffer" tool of ArcToolbox to generate circles around the locations of active network wells. Areas outside of these buffers are considered to be spatial gaps because they are beyond the maximum allowable distance for coverage to be claimed by an existing well. Gaps (areas outside of buffers) were calculated within the boundaries of each of the nine NWS climate divisions. Gap areas were then divided by 500 to determine the number of additional monitoring wells that each climate division should receive in an expansion of the network. Results are presented in Table 1.

Table 1. Spatial gaps by climate division

Climate Division	Total Extent (mi^2)	Inside Buffer (mi^2)	Outside Buffer (mi^2)	Wells Needed (Outside Buffer / 500)
Northwest	4,151.09	2,123.44	2,027.65	4
North Central	3,890.28	1,365.65	2,524.63	5
Northeast	3,585.49	1,506.48	2,079.01	4
West Central	4,120.86	1,762.30	2,358.56	5
Central	6,103.80	3,921.77	2,182.03	4
East Central	2,578.06	1,350.79	1,227.26	2
Southwest	5,019.96	2,014.41	3,005.55	6
South Central	3,945.99	510.43	3,435.56	7
Southeast	2,766.21	695.23	2,070.98	4

Optimal Network Configuration

Selecting the top n aquifer scenarios (where n was determined by the spatial gap analysis) from each of the nine climate divisions yields a list of forty-one total aquifer scenarios that need to receive new monitoring wells. Having this list, the design process then enters a final phase for deciding upon the actual geographic locations where new wells are to be installed. Being that some aquifer scenarios occur in fragmented (i.e., non-contiguous) patterns across the landscape, and that aquifer scenarios occasionally occur in the same neighborhood, a method for locating wells irrespective of geographic spacing may lead to an inefficient and clustered distribution of wells. Therefore, a spatial optimization technique known as the p-median problem (PMP) was applied to prevent clustering of proposed monitoring sites. Solving the PMP relies on a branch of mathematics known as mathematical programming that deals with problems that cannot be solved with closed form methods involving analytic techniques of algebra and geometry.

Potential Sites

There are no guidelines that dictate exactly what sorts of sites may be considered eligible for a groundwater monitoring well. Typically, wells are drilled on publicly owned land, but numerous instances may be cited of public-private partnership. Monitoring wells are not considered obnoxious as they do not generate commotion, loud noises, or foul odors, and they do not require large spaces. Thus, options for potential sites are not limited by proximity to populations that might object to the installation.

The most important consideration when evaluating a site is how closely the landscape features resemble those of the aquifer scenario for which the site is intended to represent (EPA 2002). Additional considerations include (1) promise for a site to be used over the long-term (i.e., stable ownership is desirable), (2) availability of wells from other monitoring networks, (3) distance to the nearest well from the active network, and (4) suitability of a potential site for baseline monitoring (i.e., effects of pumping should be minimal). The first consideration was honored by compiling a data set of properties having ownership unlikely to change in the foreseeable future. These included publicly managed parks and recreation areas, police and fire stations, academic institutions, and religious facilities. Second, point locations of wells from other networks were added to the list of potential sites, including wells from the NAWQA network and from localized clusters of water-quality monitoring networks in Hamilton, Lake, Porter, and St. Joseph counties (USGS 2014). Third, sites were filtered out of the list if found to be within a distance of 12.62 miles from an active network well.

The fourth consideration (baseline suitability) was not addressed directly, but an approach to site selection was taken that narrows the search down to a general locality where potential sites can then be assessed according to their likelihood of being impacted by pumping. An attempt was not made to predict drawdown or to model the radial influence of pumping from SWWFs. Furthermore, some aquifer scenarios are so heavily utilized by pumping facilities that zones of influence generated from a modeling of the radial extent of drawdown may be overly restrictive to site availability. Omitting potential monitoring sites on the presumption that they are likely impacted by pumping would have the unwanted effect of neglecting important aquifer scenarios that require monitoring, albeit under a different network objective aimed at observing aquifer response to anthropogenic stress instead of natural conditions.

Aggregation

As noted in the previous section, the approach for evaluating potential sites was to narrow the search down to general localities where one or more properties are available for consideration (Table 6). These areas are henceforth referred to as “aggregation units” because they were used as a means to aggregate many points into a smaller number of selectable units (Figure 3). By condensing points into fewer aggregation units, the computational burden of solving the p-median problem is reduced (Church 2002).

Boundaries of aggregation units were formed by the spatial intersections of civil townships and aquifer delineations. Indiana's largest township has an area of 112 square miles, which is also the maximum possible area for any of the aggregation units. However, areas of aggregation units are generally less than the total area of their constituent township because aquifer systems normally cover a smaller portion. This holds true except for instances where an entire township is underlain by a single aquifer system. Centroid points from aggregation units were used as input to the p-median problem discussed in the section that follows.

Regionally Constrained p-Median Problem

The p-median problem (PMP) belongs to a class of discrete location models residing within the field of operations research. The objective of a p-median model in its most generic form is to select locations for *p* facilities that minimize demand-weighted total distance between demand nodes and the facilities to which they are assigned (Church, ReVelle 1976). As a preliminary step to solving any PMP, real-world networks must be abstracted into conceptual graphs comprised of discrete nodes connected via some metric of distance. Optimal solutions are found by positioning medians (i.e., facilities) at nodes on the graph such that a global minimization is attained (Church, ReVelle 1976). Many practitioners have applied algorithms to solve p-median models for optimal or near-optimal solutions in complex facility location problems. While the PMP in its generic form has proved useful in numerous real-world applications, the specific design requirements of this analysis called for an adapted formulation that incorporates an additional constraint.

Categorical sampling needs were established *a priori* to the process step of locational decision-making. That is, a prioritized list of aquifer scenarios was already known before geographic examination of potential monitoring sites. An adapted formulation was needed that adheres to categorical sampling objectives while seeking a spatially optimal solution. Church (1990) introduces this adaptation by adding another constraint to the PMP that ensures a minimum/maximum number of “facilities” (i.e., wells) are assigned for each “region” (i.e., aquifer scenario) specified in the model. This modified formulation has been aptly termed the regionally constrained p-median problem (Church 1990). Notation is provided on the following page.

Minimize: $\sum_i \sum_j a_i d_{ij} x_{ij}$

$$\sum_j x_{ij} = 1$$

$$\sum_j Y_j = p$$

$$x_{ij} \leq Y_j$$

$$\sum_{j \in R_r} Y_j \leq p_r^{max}$$

$$\sum_{j \in R_r} Y_j \geq p_r^{min}$$

$$x_{ij} = 0,1$$

$$Y_j = 0,1$$

where

i = index for demand node i

j = index for facility node j

a_i = demand at node i

d_{ij} = distance from demand node i to facility node j

$$Y_j = \begin{cases} 1 & \text{if a monitoring well is located at facility node } j \\ 0 & \text{if not} \end{cases}$$

$$x_{ij} = \begin{cases} 1 & \text{if demand at node } i \text{ assigns to facility node } j \\ 0 & \text{if not} \end{cases}$$

p = total number of facilities (monitoring wells) to be located

r = index representing a region (aquifer scenario)

R_r = set of nodes comprising region r

p_r^{max} = maximum number of facilities (monitoring wells) that can be allocated to region r

p_r^{min} = minimum number of facilities (monitoring wells) that must be allocated to region r

Implementation of the RCPMP in the context of this particular application translates to locating a total of *p* new monitoring wells relative to a set of demand nodes where each region (i.e., aquifer scenario) receives exactly one well. Centroids of civil townships were chosen to represent demand nodes because they exhibit relatively uniform spacing and are spread equally throughout the state. In a groundwater monitoring network, the distribution of wells should be driven by hydrologically relevant variables, but should also be politically dispersed so as to maximize benefit for administrative subdivisions. Graph nodes designated as candidates for well placement (i.e., facility nodes) are represented by the centroids of aggregation units described earlier. Having separate sets of demand nodes and facility nodes helps reduce model complexity in cases where the number of demand nodes is smaller than the number of facility nodes. This is because the pairwise matrix (m x n) formed by distances separating demand nodes (m) and facility nodes (n) is smaller when demand nodes are a subset of the whole graph.

Euclidean distances were used as the metric for linkages between nodes in the matrix generated for this problem. Therefore, the objective of spatial optimization is simply to minimize the sum of straight-line distances that separate township centroids (i.e., demand nodes) from their nearest monitoring wells (i.e., aggregation units, facility nodes). An alternative approach might construct the graph as a hydrologic system composed of linkages that relate nodes via deterministic patterns of water movement. In order to construct such a graph at the statewide scale, regional and intermediate patterns of groundwater flow would need to be generalized throughout Indiana and at multiple depths in the vertical dimension. Large-scale simulations of dominant flow can be achieved where data are available for historical water levels and three-dimensional depictions of lithology (Arihood, Basch 1994; Zhou, Li 2011). These data can be derived from well logs that cover most of the state.

To summarize the model details that have been described, an m x n matrix was generated from separate sets of demand nodes (township centroids) and facility nodes (aggregation units); the objective function of the model minimizes the sum of Euclidean distances between these sets of nodes; a constraint is added that guarantees specific “regions” (aquifer scenarios) are assigned facility nodes (i.e., aggregation units selected to receive new monitoring wells); and the final solution set is a selection of *p* facility nodes (aggregation units) each containing a subset of potential monitoring sites.

Solving the RCPMP

The open-source programming language R was used for model formulation and solving of the RCPMP. A program was written in R to import data for facility nodes (i.e., aggregation units) and demand nodes (i.e., township centroids), execute intermediate processing, and output a file of the reference numbers that uniquely identify aggregation units included in the final solution set.

Intermediate processing involved generation of a pairwise distance matrix (m x n) storing the Euclidean (straight-line) distances that separate demand nodes (m) and facility nodes (n), which was subsequently used to model their relationships (inequalities) in an integer programming formulation of the problem. The objective function, constraints, and variable declarations that constitute the model structure were written in LP file format and a solver function from the CRAN package named lpSolveAPI was used to return an object containing the results (i.e., objective value and decisions made for variables) of the model (Konis 2014). The package lpSolveAPI uses a branch-and-bound method for solving integer programming problems (Konis 2014).

After a series of trials were run that included increasingly greater numbers of nodes in the model, it was revealed that lpSolveAPI encounters a parsing error when attempting to read an .lp file having one or more lines with excessive character length (Table 5). The objective function of the p-median model was expressed entirely on the first line of the LP-formatted text file. Thus, as more nodes were added to the model, the first line of the .lp file grew rapidly in the lateral direction; eventually becoming so long that lpSolveAPI could not interpret the file. The exact character length at which failure occurred was not determined, but it is known to lie somewhere in the range of 1,328,161 and 1,834,008 characters. These numbers correspond to LP-formatted p-median models with matrix lengths of 90,000 and 122,500, respectively.

Understanding the software limitations, a decision was made to split the problem into separate models for each of the nine climate divisions, and to solve the models in mutually exclusive runs through a loop that was built into the program. The number of nodes and iterations associated with models for each climate division are provided in Table 3. Upon completion of the loop a final solution set was pieced together by appending the results from each of the nine models into a single table.

Comparison to Random Selections

The configurations of wells generated via solution of the p-median problem are optimal and therefore superior (in terms of a minimization of the objective function) to any configurations that could be generated in a random or manual fashion. To demonstrate this, a random selection procedure was repeated for each of the climate divisions (abiding by the same constraint that pre-specified aquifer scenarios are included in the solution), and the results compared with those of the p-median solutions. Comparisons are presented in Table 2. On average, the p-median solutions were 28% more efficient than the random solutions.

Table 2. Comparison of p-median solutions to random selections

Climate Division	Sum of distances (p-median solution)	Sum of distances (random selections)	Improvement (%)
Northwest	1665.9	1919.6	13%
North Central	1263.4	1816.4	30%
Northeast	1386.7	1862.7	26%
West Central	1309.9	1621.9	19%
Central	2839.9	5047.3	44%
East Central	1227.8	1610.9	24%
Southwest	1349.7	2158.7	37%
South Central	849.1	1329.3	36%
Southeast	883.9	1175.8	25%

RESULTS

This study has produced a ranked list of aquifer scenarios (Table 7) sorted in descending order by areal extent. From this list, the top n ranked aquifer scenarios were selected from each of the nine climate divisions (for a total of forty-one) where n corresponds to the number of wells needed based on a desired well density (1/500 mi^2). The forty-one proposed monitoring wells (i.e., selected aquifer scenarios) include 14 to be installed at bedrock depth, which improve categorical coverage of bedrock extent by 45%; and 27 installed in unconsolidated aquifer systems, which improve categorical coverage of unconsolidated extent by 48%. Proposed monitoring wells inadvertently improve coverage of withdrawals by 20%. Potential monitoring sites were compiled from multiple sources and aggregated into larger units to reduce the number of nodes contained in p-median models. Optimal network configurations were generated for each climate division, and were found to be 28% more efficient (in terms of a minimized sum of distances) than random solutions. Proposed monitoring sites from the mutually exclusive p-median solutions were combined in a single statewide configuration that is not globally optimal, but effectively addresses categorical sampling gaps while ensuring that wells have maximal spacing within climate divisions.

DISCUSSION

The approach for network optimization presented in this study has some noteworthy advantages, limitations, and opportunities for additional research. The methodology used demonstrates how the statewide design problem can be abstracted as a discrete facility location model. The methods used are advantageous in that design criteria for (1) sampling a diversity of environmental variables and (2) spatial optimization can be combined in a single model.

Challenges encountered in this study arose primarily from the computational complexity of solving the PMP. Multiple strategies were taken to reduce model size and make solution of the PMP more feasible. These included use of an m x n matrix (as opposed to an n x n matrix), aggregation of potential monitoring sites, and splitting the statewide problem into nine independent models for each climate division. A lesson learned is that some degree of simplification is necessary for solving complex optimization problems.

A limitation of the methodology lies in its handling of the vertical dimension. Aquifer depths were generalized into categories of “unconsolidated” and “bedrock.” Being that planning decisions for new wells involve completion depths in addition to site locations, and that construction costs of monitoring wells derive mostly as a function of drilling depth, an improvement to this methodology might represent hydrogeological layers with more detail. Also, the piece-meal approach for assembling a statewide solution is considered a limitation. Additional research might investigate how to remedy edge effects that arise from mutually exclusive solutions for models that have adjacent spatial domains.

Lastly, groundwater is but one component of a larger hydrologic system. There is opportunity for research in exploring how to integrate wells with sensors from other hydrologic monitoring networks (e.g., stream gages, soil moisture), or similarly, how to superimpose networks with different objectives.

TABLES AND FIGURES

Table 3. Summary of active network wells

Name	Index	Rank	Well Depth	Baseline Monitoring	Continuous/Real-Time
SHELBY 2	1869	1	150	yes	Continuous
GRANT 10	503	17	198	no	Continuous
GRANT 8	1869	1	35	yes	Continuous
DECATUR 2	1869	1	43	yes	Continuous
BARTHOLOMEW 4	551	72	93	no	Real-Time
MORGAN 4	551	72	64	no	Real-Time
HAMILTON 7	551	72	82	no	Real-Time
MARION 35	917	111	83	no	Real-Time
MARION 39	917	111	28	no	Real-Time
BOONE 17	580	18	172	yes	Continuous
DELAWARE 4	512	119	91	yes	Real-Time
RANDOLPH 3	1870	5	54	yes	Real-Time
WAYNE 6	552	191	49	yes	Real-Time
FULTON 7	545	36	102	yes	Real-Time
ELKHART 4	577	99	62	no	Continuous
CASS 3	1871	2	130	yes	Real-Time
WHITLEY 3	506	53	191	yes	Continuous
WELLS 4	1872	4	79	no	Continuous
LAGRANGE 2	532	93	86	no	Real-Time
LAKE 13	867	113	23	yes	Continuous
PULASKI 7	1239	195	105	yes	Continuous
BENTON 4	515	131	310	yes	Real-Time
JASPER 13	1850	11	150	yes	Real-Time
LA PORTE 9	538	33	32	no	Real-Time
NEWTON 8	1873	6	150	no	Continuous
HARRISON 8	1832	12	93	yes	Real-Time
JEFFERSON 5	1875	27	200	yes	Continuous
CLARK 20	1289	291	100	no	Real-Time
VANDERBURGH 7	1857	3	70	yes	Real-Time
KNOX 8	1857	3	137	yes	Continuous
MARTIN 5	1867	8	143	yes	Continuous
POSEY 3	558	31	58	no	Real-Time
KNOX 7	558	31	43	no	Real-Time
TIPPECANOE 18	516	100	64	yes	Continuous
PARKE 6	1868	10	155	yes	Continuous
VIGO 7	925	178	70	no	Real-Time
MONTGOMERY 7	587	61	111	no	Continuous

Table 4. Model sizes for each climate division

Climate Division	Iterations	Facility Nodes (Aggregation Units)	Demand Nodes (Township Centroids)	Matrix Length
Northwest	3868	59	111	6549
North Central	4441	129	113	14577
Northeast	9246	197	113	22261
West Central	10034	164	114	18696
Central	9753	144	180	25920
East Central	2958	58	82	4756
Southwest	10582	259	118	30562
South Central	11995	411	92	37812
Southeast	6148	188	86	16168

Table 5. Trials to identify model size at which lpSolveAPI fails to interpret LP file

Matrix Length (m x n)	# Characters in 1st Line of Model	Parsing Error
2500	33594	N
10000	137741	N
22500	323729	N
40000	583960	N
62500	918764	N
90000	1328161	N
122500	1834008	Y

Table 6. Summary of potential monitoring sites by class

Climate Division	Index	Rank	USGS	Fire	Police	Recreation	Religious	School
Central	1836	13				1		
Central	503	17						2
Central	517	22					1	
Central	595	23					1	
East Central	1852	25					1	
East Central	581	35			1		1	
North Central	1848	7					1	
North Central	543	47	1				1	1
North Central	505	49					1	
North Central	541	62					3	
North Central	597	63				1		
Northeast	1872	4				2		
Northeast	1849	9		1	2	3	1	3
Northeast	539	30					6	1

Table 6. (cont.)

Climate Division	Index	Rank	USGS	Fire	Police	Recreation	Religious	School
Northeast	529	46						1
Northwest	1873	6	2					
Northwest	538	33				1		
Northwest	534	43				3		
Northwest	1838	51					1	
South Central	1839	16				1		
South Central	1132	19				1		
South Central	1842	26		1	1	2	6	1
South Central	1376	29				1		
South Central	1254	37					2	1
South Central	2399	60					7	
South Central	522	64				1	1	1
Southeast	1856	14		1	1	2	5	1
Southeast	1255	32			1	1		
Southeast	523	38				1	1	
Southeast	1133	44					1	
Southwest	524	15					3	
Southwest	1846	20					1	
Southwest	558	31	2					
Southwest	494	41					1	
Southwest	1256	42					2	
Southwest	1378	54				1		
West Central	1841	21					1	
West Central	603	24		1		1		1
West Central	525	28						1
West Central	1835	45					1	
West Central	1847	50					5	

Figure 2. Map of active network wells

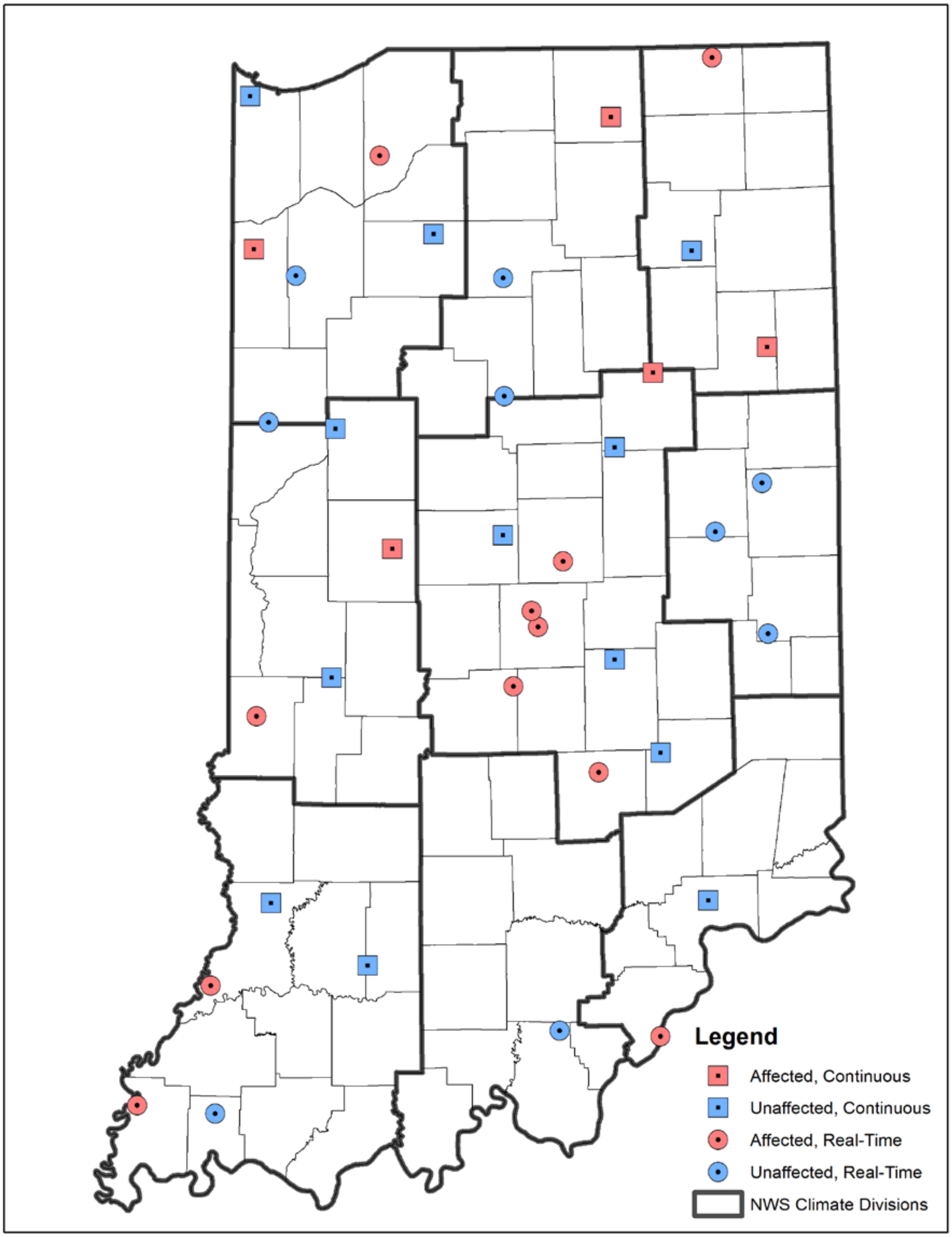


Figure 3. Map of potential monitoring sites within an aggregation unit

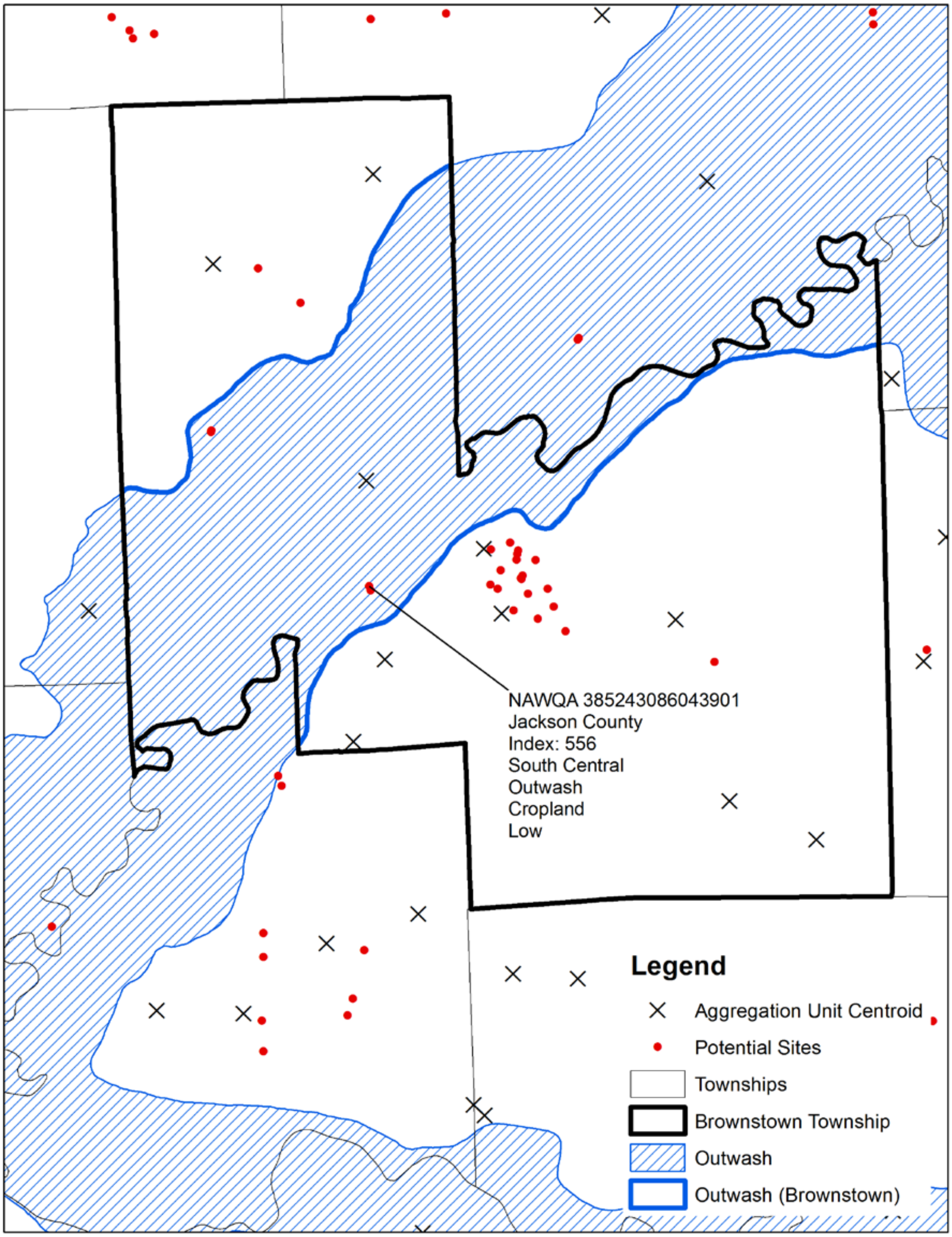


Figure 4. Map of spatial gaps under active network

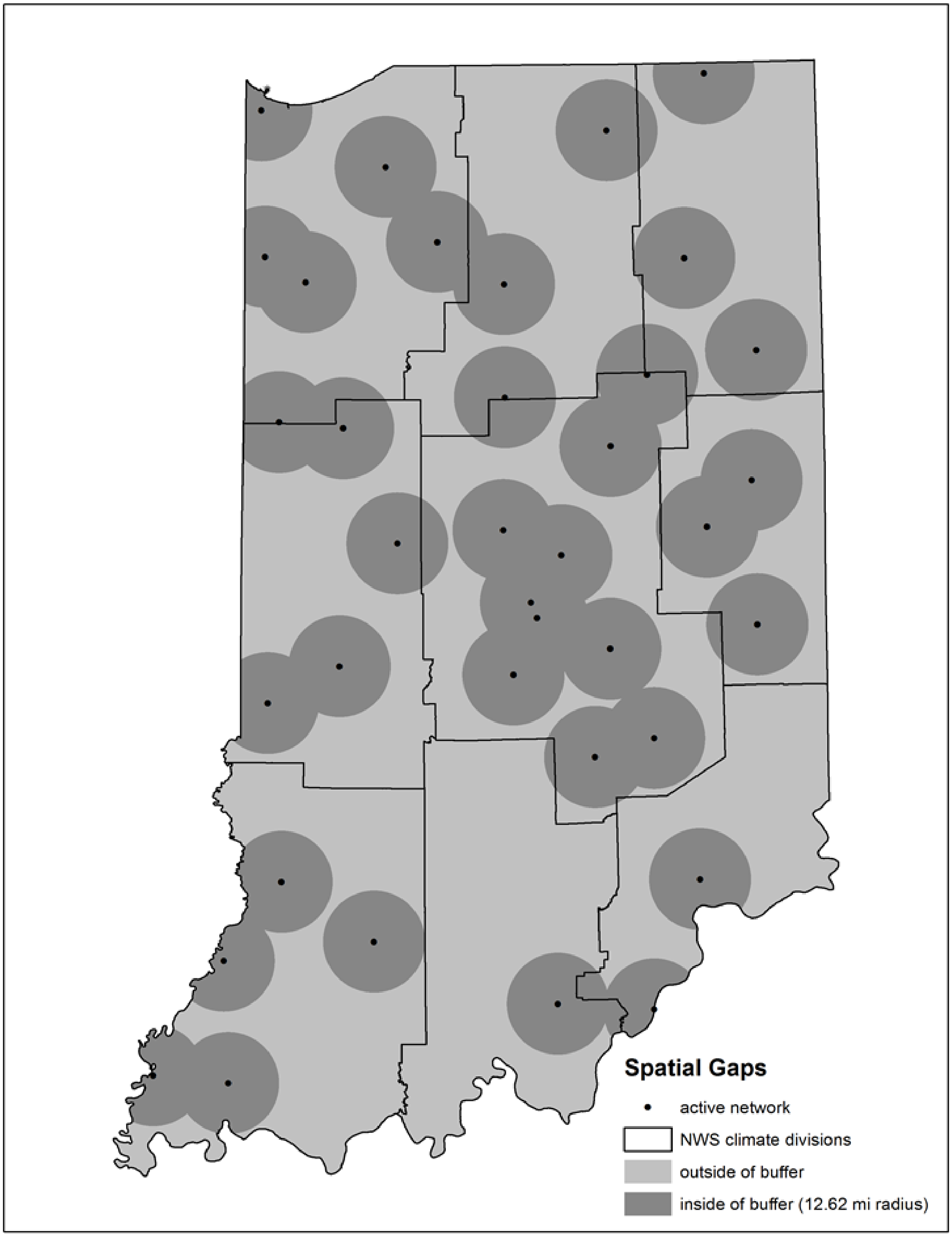


Figure 5. Map showing configuration of proposed monitoring sites

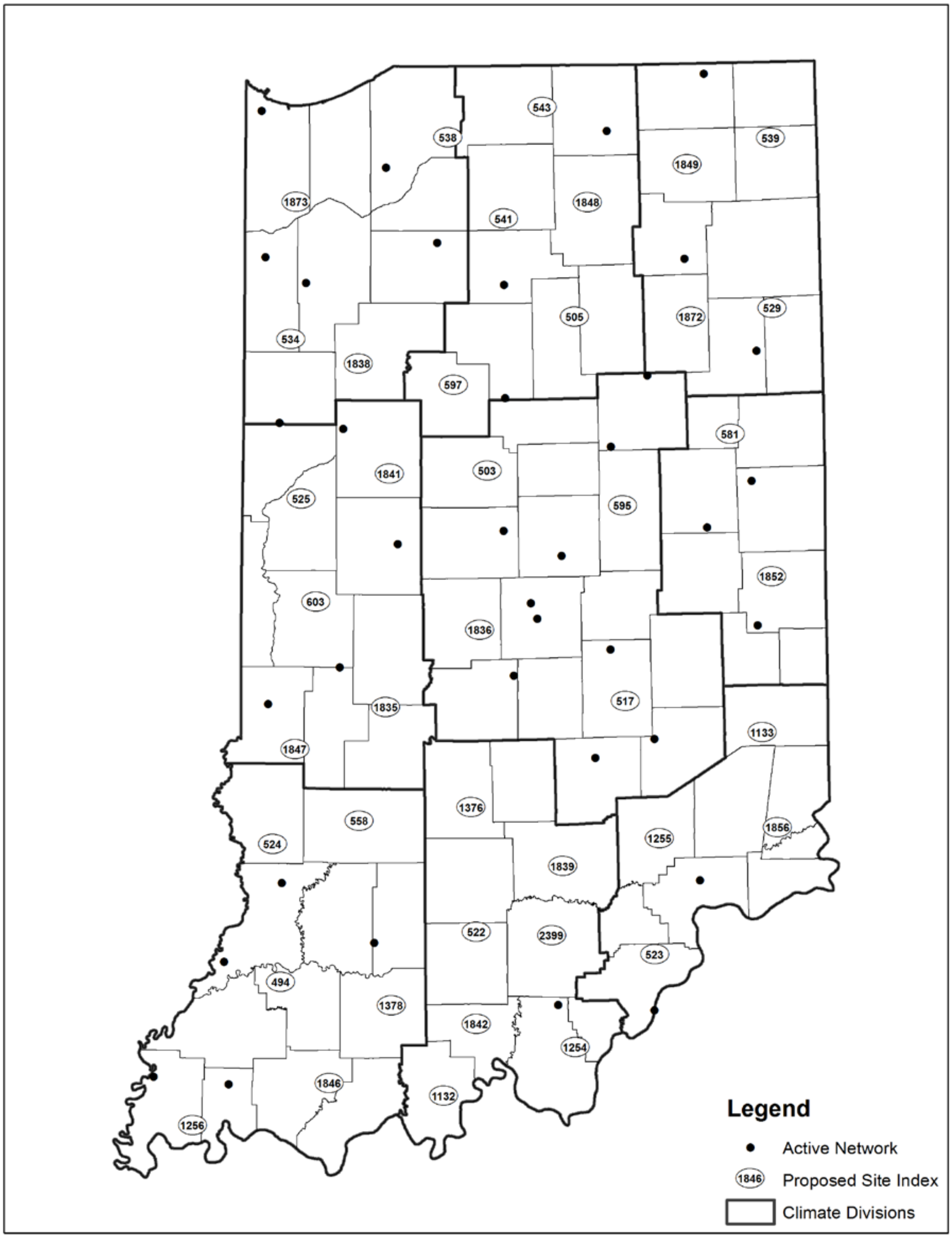


Figure 6. Map of nodes used in p-median models

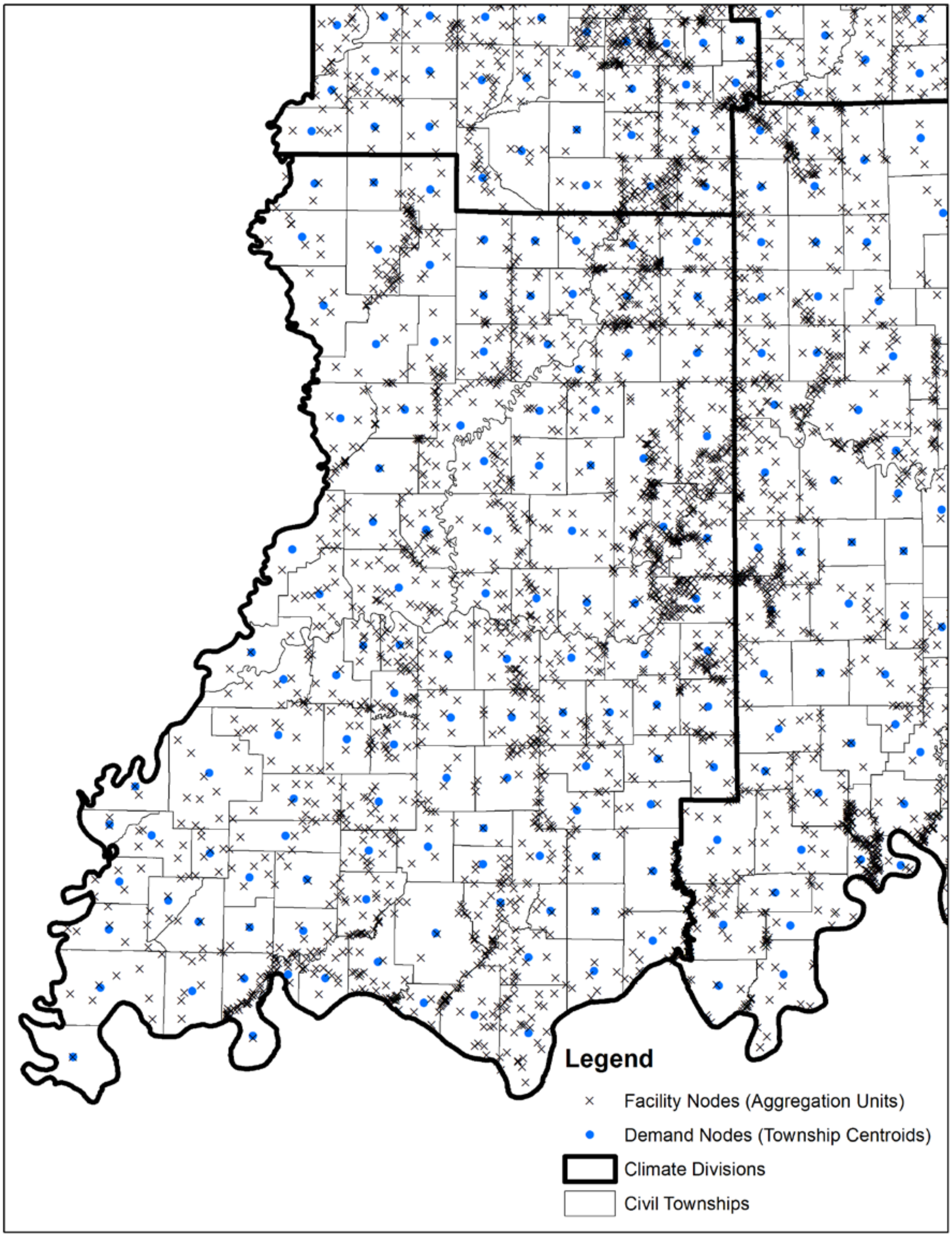


Figure 7. Spatial data sets of environmental variables used in stratification scheme

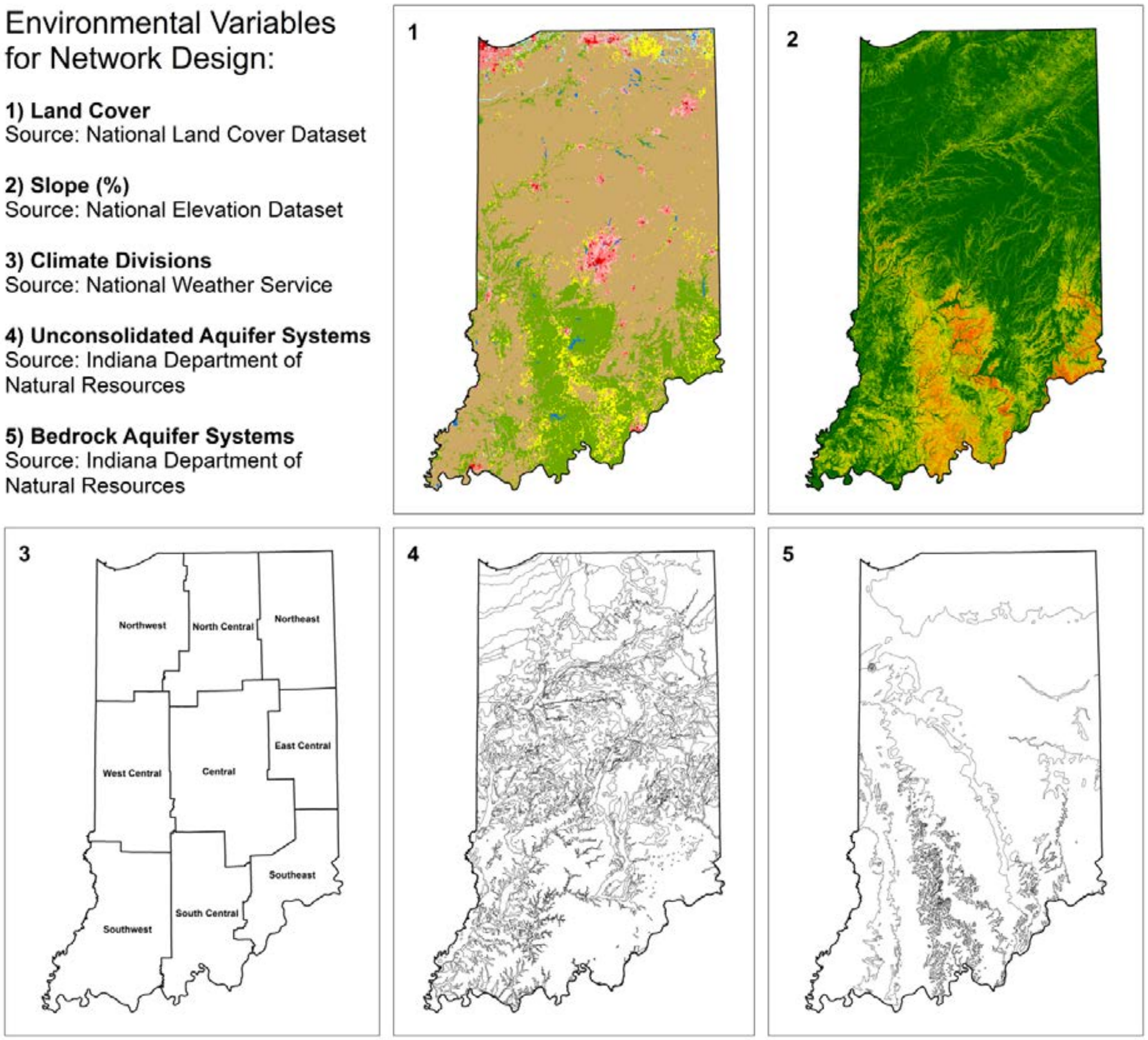


Table 7. Ranked list of aquifer scenarios

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
1869	Bedrock	Central	Silurian and Devonian Carbonates	NA	NA	3		3942.06	6942.15	1	227
1871	Bedrock	North Central	Silurian and Devonian Carbonates	NA	NA	1		2165.38	1397.28	2	230
1857	Bedrock	Southwest	McLeansboro Group	NA	NA	2		1952.53	27.33	3	149
1872	Bedrock	Northeast	Silurian and Devonian Carbonates	NA	NA		1	1951.57	4623.26	4	288
1870	Bedrock	East Central	Silurian and Devonian Carbonates	NA	NA	1		1733.37	2669.44	5	192
1873	Bedrock	Northwest	Silurian and Devonian Carbonates	NA	NA		1	1699.62	5680.92	6	278
1848	Bedrock	North Central	Coldwater, Ellsworth, and Antrim Shales	NA	NA			1657.87	30.34	7	115
1867	Bedrock	Southwest	Raccoon Creek Group	NA	NA	1		1648.3	30.01	8	102
1849	Bedrock	Northeast	Coldwater, Ellsworth, and Antrim Shales	NA	NA			1628.44	143.38	9	159
1868	Bedrock	West Central	Raccoon Creek Group	NA	NA	1		1598.7	39.89	10	207
1850	Bedrock	Northwest	Coldwater, Ellsworth, and Antrim Shales	NA	NA	1		1572.57	224.85	11	162
1832	Bedrock	South Central	Blue River and Sanders Groups	NA	NA	1		1481.12	143.18	12	147
1836	Bedrock	Central	Borden Group	NA	NA			1427.69	110.62	13	116
1856	Bedrock	Southeast	Maquoketa Group	NA	NA			1327.99	3.43	14	150
524	Unconsolidated	Southwest	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			1308.49	60.1	15	75
1839	Bedrock	South Central	Borden Group	NA	NA			1287.34		16	NA
503	Unconsolidated	Central	Complex	Cultivated Crops	Low		1	1207.18	4863.72	17	123
580	Unconsolidated	Central	Till	Cultivated Crops	Low	1		1176.23	976.79	18	111
1132	Unconsolidated	South Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	High			1143.59		19	NA

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
1846	Bedrock	Southwest	Carbondale Group	NA	NA			1120.66		20	168
1841	Bedrock	West Central	Borden Group	NA	NA			1012.14	79.96	21	168
517	Unconsolidated	Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			1009.36	91.2	22	60
595	Unconsolidated	Central	Till Subsystem	Cultivated Crops	Low			899.16	639.23	23	100
603	Unconsolidated	West Central	Till Subsystem	Cultivated Crops	Low			893.69	63.91	24	121
1852	Bedrock	East Central	Maquoketa Group	NA	NA			844.45	14	25	30
1842	Bedrock	South Central	Buffalo Wallow, Stephensport, and West Baden Groups	NA	NA			812.68		26	NA
1875	Bedrock	Southeast	Silurian and Devonian Carbonates	NA	NA	1		780.93	7.82	27	16
525	Unconsolidated	West Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			742.04	44.1	28	62
1376	Unconsolidated	South Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Moderate			716.48		29	NA
539	Unconsolidated	Northeast	Kendallville	Cultivated Crops	Low			687.73	1605.27	30	143
558	Unconsolidated	Southwest	Outwash	Cultivated Crops	Low		2	668.35	22612.69	31	64
1255	Unconsolidated	Southeast	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Low			652.62	59.42	32	44
538	Unconsolidated	Northwest	Kankakee	Cultivated Crops	Low		1	631.84	5274.77	33	46
1859	Bedrock	Central	New Albany Shale	NA	NA			616.32	23.6	34	109
581	Unconsolidated	East Central	Till	Cultivated Crops	Low			603.9	60.34	35	113

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
545	Unconsolidated	North Central	Natural Lakes and Moraines	Cultivated Crops	Low	1		603.07	3651.29	36	133
1254	Unconsolidated	South Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Low			597.6		37	NA
523	Unconsolidated	Southeast	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			591.91	11.84	38	65
596	Unconsolidated	East Central	Till Subsystem	Cultivated Crops	Low			559.93	34.79	39	91
504	Unconsolidated	East Central	Complex	Cultivated Crops	Low			549.73	597.76	40	102
494	Unconsolidated	Southwest	Alluvial, Lacustrine, and Backwater Deposits	Cultivated Crops	Low			509.23	5.29	41	66
1256	Unconsolidated	Southwest	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Low			503.89		42	NA
534	Unconsolidated	Northwest	Iroquois Basin	Cultivated Crops	Low			491.12	61.95	43	88
1133	Unconsolidated	Southeast	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	High			490.67	719.46	44	86
1835	Bedrock	West Central	Blue River and Sanders Groups	NA	NA			474.6	39.5	45	218
529	Unconsolidated	Northeast	Hessen Cassel	Cultivated Crops	Low			469.73	31.54	46	172
543	Unconsolidated	North Central	Nappanee	Cultivated Crops	Low			454.21	1481.16	47	127
1863	Bedrock	Southeast	New Albany Shale	NA	NA			451.99		48	NA
505	Unconsolidated	North Central	Complex	Cultivated Crops	Low			446.3	2131.17	49	109
1847	Bedrock	West Central	Carbondale Group	NA	NA			442.31	0.1	50	106
1838	Bedrock	Northwest	Borden Group	NA	NA			436.89	21.81	51	334

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
1861	Bedrock	Northwest	New Albany Shale	NA	NA			426.14	326.6	52	183
506	Unconsolidated	Northeast	Complex	Cultivated Crops	Low	1		422.55	593.89	53	182
1378	Unconsolidated	Southwest	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Moderate			421.35		54	NA
555	Unconsolidated	Northwest	Outwash	Cultivated Crops	Low			419.26	682.55	55	75
1377	Unconsolidated	Southeast	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Moderate			391.52	46.52	56	70
1134	Unconsolidated	Southwest	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	High			379.43		57	35
507	Unconsolidated	Northwest	Complex	Cultivated Crops	Low			376.81	2653.16	58	111
1257	Unconsolidated	West Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Low			362.6	14.85	59	37
2399	Unconsolidated	South Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Pasture/Hay	Low			349.43		60	NA
587	Unconsolidated	West Central	Till	Cultivated Crops	Low		1	322.38	559.2	61	123
541	Unconsolidated	North Central	Maxinkuckee Moraine	Cultivated Crops	Low			318	2378.8	62	109
597	Unconsolidated	North Central	Till Subsystem	Cultivated Crops	Low			312.68	2.8	63	98
522	Unconsolidated	South Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			302.38		64	NA
582	Unconsolidated	North Central	Till	Cultivated Crops	Low			292.51	2.6	65	143
520	Unconsolidated	Northeast	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			289.35	15.6	66	207

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
1864	Bedrock	West Central	New Albany Shale	NA	NA			284.89	11.7	67	120
1844	Bedrock	Southwest	Buffalo Wallow, Stephensport, and West Baden Groups	NA	NA			275.42	25.79	68	158
946	Unconsolidated	Central	Till	Developed	Low			272.69	1264.15	69	106
610	Unconsolidated	Northwest	Valparaiso Moraine	Cultivated Crops	Low			253.08	377.39	70	98
1866	Bedrock	South Central	Raccoon Creek Group	NA	NA			245.14		71	496
551	Unconsolidated	Central	Outwash	Cultivated Crops	Low		3	242.65	2153.9	72	76
584	Unconsolidated	Northwest	Till	Cultivated Crops	Low			225.21	15.6	73	115
583	Unconsolidated	Northeast	Till	Cultivated Crops	Low			221.78	11.6	74	109
553	Unconsolidated	North Central	Outwash	Cultivated Crops	Low			221.52	5762.26	75	102
646	Unconsolidated	Southwest	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Moderate			216.79		76	NA
598	Unconsolidated	Northeast	Till Subsystem	Cultivated Crops	Low			215.72		77	NA
599	Unconsolidated	Northwest	Till Subsystem	Cultivated Crops	Low			212.37	61.83	78	85
1379	Unconsolidated	West Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Moderate			207.14		79	NA
536	Unconsolidated	Northwest	Iroquois Moraine	Cultivated Crops	Low			205.92	16.82	80	81
2400	Unconsolidated	Southeast	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Pasture/Hay	Low			197.63		81	NA
559	Unconsolidated	West Central	Outwash	Cultivated Crops	Low			188.99	6668.3	82	94
518	Unconsolidated	East Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			184.1		83	NA

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
1858	Bedrock	West Central	McLeansboro Group	NA	NA			183.26		84	NA
521	Unconsolidated	Northwest	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			182.66	26.86	85	41
546	Unconsolidated	Northeast	Natural Lakes and Moraines	Cultivated Crops	Low			177.31	1868.52	86	121
1840	Bedrock	Southeast	Borden Group	NA	NA			176.86		87	NA
2521	Unconsolidated	South Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Pasture/Hay	Moderate			176.73		88	NA
1335	Unconsolidated	West Central	Till Subsystem	Forest	Low			171.86	0.2	89	122
943	Unconsolidated	North Central	St. Joseph Aquifer System and Tributary Valleys Sole Source Aquifer	Developed	Low			171.57	16576.35	90	98
556	Unconsolidated	South Central	Outwash	Cultivated Crops	Low			150.65	438.21	91	69
570	Unconsolidated	Southwest	Outwash Subsystem	Cultivated Crops	Low			150.01	466.16	92	63
532	Unconsolidated	Northeast	Howe Outwash	Cultivated Crops	Low		1	148.12	6157.62	93	110
519	Unconsolidated	North Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Cultivated Crops	Low			146.9		94	NA
2401	Unconsolidated	Southwest	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Pasture/Hay	Low			137.89		95	NA
508	Unconsolidated	West Central	Complex	Cultivated Crops	Low			135.85	645.82	96	126
1135	Unconsolidated	West Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	High			135.39	7.8	97	85
528	Unconsolidated	Northwest	Eolian Sands	Cultivated Crops	Low			133.07	143.46	98	100

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
577	Unconsolidated	North Central	St. Joseph Aquifer System and Tributary Valleys Sole Source Aquifer	Cultivated Crops	Low		1	128.73	4695.64	99	93
516	Unconsolidated	West Central	Complex (over buried valley with some potential)	Cultivated Crops	Low	1		121.61	113.67	100	184
1249	Unconsolidated	Central	Dissected Till and Residuum / Till Veneer / Unglaciated Southern Hills and Lowlands (southern IN only)	Forest	Low			117.07		101	NA
1862	Bedrock	South Central	New Albany Shale	NA	NA			112.93	3	102	65
1226	Unconsolidated	Southwest	Alluvial, Lacustrine, and Backwater Deposits	Forest	Low			112.72		103	NA
571	Unconsolidated	West Central	Outwash Subsystem	Cultivated Crops	Low			110.36	865.7	104	74
869	Unconsolidated	Central	Complex	Developed	Low			101.67	5756.55	105	121
1851	Bedrock	Central	Maquoketa Group	NA	NA			100.73		106	NA
2416	Unconsolidated	Northeast	Kendallville	Pasture/Hay	Low			99.48	2.64	107	127
917	Unconsolidated	Central	Outwash	Developed	Low		2	89.08	16554.79	111	74
867	Unconsolidated	Northwest	Calumet	Developed	Low	1		87.53	2266.57	113	28
510	Unconsolidated	East Central	Complex (over buried valley with no or unknown potential)	Cultivated Crops	Low			86.75	693.82	114	119
976	Unconsolidated	Northwest	Valparaiso Moraine	Developed	Low			79.75	1483.14	118	129
512	Unconsolidated	East Central	Complex (over buried valley with some potential)	Cultivated Crops	Low	1		72.97	56.97	119	106
537	Unconsolidated	North Central	Kankakee	Cultivated Crops	Low			70.36	3944.97	121	84
515	Unconsolidated	Northwest	Complex (over buried valley with some potential)	Cultivated Crops	Low	1		60.54	37.26	131	140
905	Unconsolidated	Northeast	Kendallville	Developed	Low			54.22	2935.78	134	159
513	Unconsolidated	North Central	Complex (over buried valley with some potential)	Cultivated Crops	Low			48.01	719.99	139	196
3141	Unconsolidated	Northeast	Howe Outwash	Wetlands	Low			35.57	1122.71	159	85

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
1291	Unconsolidated	West Central	Outwash	Forest	Low			26.06	6656.51	176	85
925	Unconsolidated	West Central	Outwash	Developed	Low		1	25.85	3143.44	178	93
1288	Unconsolidated	South Central	Outwash	Forest	Low			25.47	811.91	180	101
882	Unconsolidated	West Central	Complex (over buried valley with some potential)	Developed	Low			24.73	1974.05	183	203
552	Unconsolidated	East Central	Outwash	Cultivated Crops	Low	1		23.37	248.24	191	79
1239	Unconsolidated	Northwest	Complex	Forest	Low	1		21.51	9.96	195	89
511	Unconsolidated	Central	Complex (over buried valley with some potential)	Cultivated Crops	Low			21.44	1035.12	196	197
548	Unconsolidated	North Central	Natural Lakes and Moraines Subsystem	Cultivated Crops	Low			21.14	1544.65	198	129
2069	Unconsolidated	Southwest	Outwash	Open Water	Low			20.74	4233.89	200	95
578	Unconsolidated	Northeast	St. Joseph Aquifer System and Tributary Valleys Sole Source Aquifer	Cultivated Crops	Low			20.49	1178.39	201	91
1283	Unconsolidated	Central	Outwash	Forest	Low			20.25	2376.45	205	93
919	Unconsolidated	North Central	Outwash	Developed	Low			18.53	2237.8	219	111
557	Unconsolidated	Southeast	Outwash	Cultivated Crops	Low			17.76	2503.47	221	100
608	Unconsolidated	Northeast	Topeka	Cultivated Crops	Low			17.65	1065.38	223	147
909	Unconsolidated	North Central	Nappanee	Developed	Low			13.67	699.3	247	117
1289	Unconsolidated	Southeast	Outwash	Forest	Low		1	9.69	1533.16	291	100
874	Unconsolidated	West Central	Complex	Developed	Low			7.47	1676.19	327	122
2434	Unconsolidated	Southeast	Outwash	Pasture/Hay	Low			5.42	1486.42	366	105
923	Unconsolidated	Southeast	Outwash	Developed	Low			5.02	5264.67	377	92

Table 7. (cont.)

Index[1]	Setting	Climate Division	Aquifer System	Majority Land Cover	Topo-graphy	Baseline Wells	Affected Wells	Areal Extent[2]	Annual Withdrawals[3]	Rank[4]	Average Pumping Depth[5]
1039	Unconsolidated	Central	Outwash	Developed	Moderate			4.59	732.51	387	66
2436	Unconsolidated	West Central	Outwash	Pasture/Hay	Low			4.32	1464.36	398	98
673	Unconsolidated	Central	Outwash	Cultivated Crops	Moderate			3.88	917.96	419	95
1294	Unconsolidated	West Central	Outwash (over buried valley with some potential)	Forest	Low			2.47	1195.36	499	105
1284	Unconsolidated	East Central	Outwash	Forest	Low			2.3	771.41	518	93
1047	Unconsolidated	West Central	Outwash	Developed	Moderate			0.88	1057.75	703	102
…	…	…	…	…	…	…	…	…	…	…	…

[1] unique index assigned for aquifer scenarios (142 of the 3,341 aquifer scenarios are listed in Table 7)
[2] total areal extent of Indiana: 36,156 square miles
[3] total annual withdrawals for all aquifer scenarios in Indiana: 261,850 (millions of gallons)
[4] rank numbers assigned by descending order of areal extent
[5] average pumping depth (units of feet) calculated by dividing the sum of SWWF pumping depths by the total number of SWWFs in an aquifer scenario

REFERENCES

Articles

Arihood, Leslie D., and Mark Edward Basch. "Geohydrology and simulated ground-water flow in an irrigated area of northwestern Indiana." (1994). Church, Richard, and Charles R. Velle. "The maximal covering location problem." *Papers in regional science* 32, no. 1 (1974): 101-118.

Church, Richard L. "The Regionally Constrained p-Median Problem."*Geographical Analysis* 22, no. 1 (1990): 22-32.

Church, Richard L. "Geographical information systems and location science."*Computers & Operations Research* 29, no. 6 (2002): 541-562.

Church, Richard L., and Charles S. ReVelle. "Theoretical and Computational Links between the p-Median, Location Set-covering, and the Maximal Covering Location Problem." *Geographical Analysis* 8, no. 4 (1976): 406-415.

Fisher, Jason C. "Optimization of Water-Level Monitoring Networks in the Eastern Snake River Plain Aquifer Using a Kriging-Based Genetic Algorithm Method." In *AGU Fall Meeting Abstracts*, vol. 1, p. 04. 2013.

Heath, Ralph C. "Design of Ground–Water Level Observation–Well Programs."*Groundwater* 14, no. 2 (1976): 71-77.

Hudak, Paul F., and Hugo A. Loaiciga. "A location modeling approach for groundwater monitoring network augmentation." *Water Resources Research* 28, no. 3 (1992): 643-649.

Konis, Kjell. LpSolveAPI Version 5.5.2.0-12. September 2, 2014. Accessed November 6, 2014. http://cran.r-project.org/web/packages/lpSolveAPI/lpSolveAPI.pdf.

Nabi, Aadil, Adrian H. Gallardo, and Shakeel Ahmed. "Optimization of a groundwater monitoring network for a sustainable development of the Maheshwaram Catchment, India." *Sustainability* 3, no. 2 (2011): 396-409.

Nolan, Bernard T. "Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States." *Groundwater* 39, no. 2 (2001): 290-299.

Pearson, Scott, Jan Falteisek, and Jim Berg. "Minnesota Groundwater Level Monitoring Network - Guidance Document for Network Development." Groundwater Level Monitoring Program. May 1, 2011. Accessed November 5, 2014. http://files.dnr.state.mn.us/publications/waters/groundwater_network_guidance.pdf.

"Purposive Sampling." Laerd. Accessed November 9, 2014. http://dissertation.laerd.com/purposive-sampling.php#maximum-variation-sampling.

"Stratified Sampling." Wikipedia. Accessed November 9, 2014. http://en.wikipedia.org/wiki/Stratified_sampling.

U.S. Environmental Protection Agency. Office of Environmental Information. *Guidance on Choosing a Sampling Design for Environmental Data Collection*. EPA QA/G-5S. Washington, D.C.: U.S. United States Government Printing Office, 2002.

Wittman, Jack. "Modernizing the State's Approach to a Critical Resource." Water and Economic Development in Indiana. August 8, 2014. Accessed November 6, 2014. http://www.indianachamber.com/index.php/water-study.

Zhou, Wanfang, Barry F. Beck, Arthur J. Pettit, and Brad J. Stephenson. "A groundwater tracing investigation as an aid of locating groundwater monitoring stations on the Mitchell Plain of southern Indiana." *Environmental Geology* 41, no. 7 (2002): 842-851.

Zhou, Yangxiao, and Wenpeng Li. "A review of regional groundwater flow modeling." *Geoscience Frontiers* 2, no. 2 (2011): 205-214.

Data sets

Bernardin, Lochmueller & Associates (2007). Religious Centers in Indiana, derived from Geographic Names Information System. [ESRI shapefile] Bloomington, IN: Indiana Geological Survey [distributor] 11 Nov. 2014. http://maps.indiana.edu/metadata/Infrastructure/Facilities_Religous_Facilities.html

Indiana Department of Natural Resources (2011). Bedrock Aquifer Systems of Indiana (1:500,000). [ESRI shapefile] Bloomington, IN: Indiana Geological Survey [distributor]. 11 Nov. 2014. http://maps.indiana.edu/metadata/Hydrology/Aquifer_Systems_Bedrock_IDNR.html

Indiana Department of Natural Resources (2009). Outdoor Recreational Facilities in Indiana. [ESRI shapefile] Bloomington, IN: Indiana Geological Survey [distributor] 11 Nov. 2014. http://maps.indiana.edu/metadata/Infrastructure/Recreation_Facilities_IDNR.html

Indiana Department of Natural Resources (2014). Significant Water Withdrawal Facility Data [.xls spreadsheet] Indianapolis, IN: IDNR, Division of Water [distributor] 11 Nov. 2014. http://www.in.gov/dnr/water/files/metadata-wu.pdf

Indiana Department of Natural Resources (2011). Unconsolidated Aquifer Systems of Indiana (1:48,000). [ESRI shapefile] Bloomington, IN: Indiana Geological Survey [distributor]. 11 Nov. 2014. http://maps.indiana.edu/metadata/Hydrology/Aquifer_Systems_Unconsolidated_IDNR.html

The Polis Center (2011). Fire Stations in Indiana, derived from Essential Facilities Data of the Multi-Hazard Mitigation Planning Data. [ESRI shapefile] Bloomington, IN: Indiana Geological Survey [distributor]. 11 Nov. 2014. http://maps.indiana.edu/metadata/Infrastructure/Critical_Facilities_MHMP_Fire_Stations.html

The Polis Center (2011). Police Stations in Indiana, derived from Essential Facilities Data of the Multi-Hazard Mitigation Planning Data. [ESRI shapefile] Bloomington, IN: Indiana Geological Survey [distributor]. 11 Nov. 2014. http://maps.indiana.edu/metadata/Infrastructure/Critical_Facilities_MHMP_Police_Stations.html

The Polis Center (2011). School Facilities in Indiana, derived from Essential Facilities Data of the Multi-Hazard Mitigation Planning Data. [ESRI shapefile] Bloomington, IN: Indiana Geological Survey [distributor]. 11 Nov. 2014. http://maps.indiana.edu/metadata/Infrastructure/Critical_Facilities_MHMP_Schools.html

U.S. Department of Commerce (2014). TIGER/Line Shapefile, rd13, 2010 state, Indiana, 2010 Census County and Equivalent State-based. [ESRI shapefile] Washington, DC: U.S. Census Bureau, Geography Division, Geographic Products Branch [distributor]. 11 Nov. 2014.

U.S. Department of Commerce (2014). TIGER/Line Shapefile, rd13, 2010 state, Indiana, 2010 Census County Subdivision State-based. [ESRI shapefile] Washington, DC: U.S. Census Bureau, Geography Division, Geographic Products Branch [distributor]. 11 Nov. 2014.

U.S. Geological Survey (2014). Groundwater Watch: Indiana Active Water Level Network. [online] 11 Nov. 2014. http://groundwaterwatch.usgs.gov/StateMap.asp?sa=IN&sc=18.

U.S. Geological Survey (2006). Land Cover in Indiana, Derived from the 2006 Land Cover Database. [30-meter TIFF Image] Bloomington, IN: Indiana Geological Survey [distributor]. 11 Nov. 2014. http://maps.indiana.edu/metadata/Environment/Land_Cover_2006.html

U.S. Geological Survey (1999). National Elevation Dataset (1 arc-second). [ArcGrid]. Sioux Falls, SD: EROS Data Center. 11 Nov 2014. http://ned.usgs.gov/ned.html

CURRICULUM VITAE

Benjamin J Sperl

EDUCATION

MS, Geographic Information Science, 2014 – Indiana University, Indianapolis, IN

Graduate Certificate, Emergency Management, 2014 – Indiana University, Indianapolis, IN

BA, Geography, 2011 – Indiana University, Bloomington, IN

PROFESSIONAL

Internships at:

U.S. Geological Survey, Indiana Water Science Center (May 2014 – Dec. 2014)

The Polis Center (Dec. 2012 – May 2013; Sep. 2013 – May 2014)

Indiana Department of Transportation, Land & Aerial Survey Office (May 2013 – Aug. 2013)

Indiana Geological Survey, Center for Geospatial Data Analysis (Feb. 2012 – Dec. 2012)

Indiana University, Center for the Study of Institutions, Populations, and Environmental Change (May 2011 – Aug. 2011)

Indiana Department of Natural Resources, Division of Parks and Reservoirs (May 2011 – Nov. 2011)